THE ASTROPHYSICAL JOURNAL, 433:479–493, 1994 October 1

THE DYNAMICAL EVOLUTION OF POOR CLUSTERS OF GALAXIES: GROWTH AND PROPERTIES OF THE FIRST-RANKED GALAXY

PAUL W. BODE,[1] ROBERT C. BERRINGTON, HALDAN N. COHN, AND PHYLLIS M. LUGGER
Astronomy Department, Indiana University, Bloomington, IN 47405
Received 1994 January 26; accepted 1994 April 7

ABSTRACT

We report N-body simulations of the dynamical evolution of isolated clusters of 50 galaxies containing a dark matter component that comprises 90% of the cluster mass. For our adopted physical scaling, the line-of-sight velocity dispersion of the cluster is 310 km s^{-1} and the initial core radius is 250 kpc. Our results are applicable to (1) present-day poor clusters, (2) the small systems that may have merged to produce present-day rich clusters, and (3) virialized subclumps within larger systems, in between major substructure merger events.

We have evolved a total of 10 cluster models, using $N = 40{,}000$ particles per model. The models are fully self-consistent in that each galaxy is represented as an extended structure containing many particles and the gravitational potential arises from the particles alone. Dark matter is apportioned between the galaxy halos and a smoothly distributed common group halo, the intracluster background (ICB). The percentage of cluster mass initially in the ICB, β, is chosen to be 50, 75, or 90. Increasing β has the effect of removing mass from dark halos around galaxies and distributing it throughout the cluster. The initial conditions were constructed by randomly sampling a King distribution with $W_0 = 6$. The galaxies are also King models; the masses of the galaxies follow a Schechter distribution function.

The five $\beta = 50$ models all followed a similar pattern of behavior. Galaxies experience dynamical friction and undergo orbital decay, leading to an enhanced encounter rate. In $\sim$10 Gyr, merging has resulted in the formation of a dominant, centrally located galaxy. Almost all of the subsequent merging involves this dominant galaxy accreting the others. Mass segregation is apparent, leading the largest galaxies to preferentially engage in merging. Merging produces an extension of the galaxy mass distribution to higher masses, while at the same time it reduces the characteristic mass of the distribution owing to the overall depletion of bright galaxies.

Once the first-ranked galaxy (FRG) has grown to twice the size of the initially largest galaxy, its velocity has typically decreased to less than half the cluster velocity dispersion and it remains within the cluster core. The distribution of FRG peculiar velocities at this point contains no values greater than the cluster dispersion; there are no high-velocity FRGs of the sort that have been observed in $\sim$10% of clusters.

The most evident change in the cluster space density profile occurs in the inner 200 kpc, where a rise in density causes the core to be erased. If the location of the FRG is taken to define the cluster center, then the density profile is even more strongly cusped and resembles a singular isothermal sphere. The FRG-centered surface density profile can be fit by both power-law and exponential profiles.

Once the FRG has assumed a central position in the cluster, multiple nuclei are seen at least 20% of the time, roughly what is expected from the projected surface density distribution. The frequency rises above this to $\sim$40% at $\sim$11 Gyr. The additional nuclei are on orbits which bring them into contact with the FRG. After these satellites merge with the FRG, the frequency of multiple nuclei falls back to the value expected from projection.

Observations of ΔM_{12}, FRG luminosity, and the number of multiple nuclei can best be fit by cluster models with ages $\sim$11 Gyr; growth in luminosity of the FRG during this amount of time is consistent with only weak cannibalism. Fitting observations of the peculiar velocities of the FRG requires younger ages of $\sim$8 Gyr.

Increasing β to 75 slows the rate of merging, but otherwise causes little change in behavior. For $\beta = 90$, the onset of merging can be delayed for over 13 Gyr; thus a dominant central galaxy is not created.

Subject headings: galaxies: clustering — galaxies: interactions — intergalactic medium — methods: numerical

1. INTRODUCTION

1.1. *Overview*

Despite intensive observational and theoretical study of clusters of galaxies, there are many unresolved questions concerning the effect of dynamical evolution on clusters and their constituent galaxies. Many dynamical processes operate under virial equilibrium conditions. These include interactions between galaxies and the dark-matter component of the cluster via dynamical friction and galaxy-galaxy interactions via two-body relaxation, merging, and tidal stripping. If there is significant substructure in clusters, then the overall cluster potential will evolve on a dynamical timescale and mean-field relaxation will also be an important process. The properties of the first-ranked galaxy (FRG) in a cluster may be affected by these processes; for recent reviews see Lugger (1991), Richstone (1990), and Merritt (1988). Simulations of the dynamical interaction of galaxies in clusters have produced conflicting results, as discussed below. We undertook the present study to reinvestigate this problem by taking advantage of the enormous

[1] bode@pegasus2.indiana.edu.

increase in available computational power that has occurred during the past decade.

1.2. *Theoretical Background*

The observation that some clusters contain a central, giant cD galaxy has led to the theory that this galaxy achieved its present size by "cannibalizing" other galaxies over the life of the cluster. The original prediction of galactic cannibalism rates in rich clusters (Ostriker & Tremaine 1975; Hausman & Ostriker 1978) is based on a highly approximate theory that uses average interaction rates for galaxies in a cluster core without considering actual galaxy trajectories; these interaction rates were sampled using Monte Carlo techniques in the latter study. These studies predicted high galactic cannibalism rates in rich ($N \sim 10^3$), high-velocity-dispersion clusters ($\sigma \sim 10^3$ km s^{-1}), with the FRG growing by $>10L^*$ over the cluster lifetime.

The more detailed simulations by Richstone & Malumuth (1983) and Malumuth & Richstone (1984) followed the trajectories of individual galaxies moving in a fixed background containing 85% of the cluster mass, while still using a Monte Carlo treatment of galaxy-galaxy interactions. Internal galactic structure was not resolved; instead, cross sections derived from other simulations of galaxy collisions were used to model merging and tidal stripping. Roughly 25% of a set of different realizations of the same initial conditions were found to produce a cD galaxy, independent of cluster richness. This suggests that statistical fluctuations play an important role in determining the morphological types of clusters. Stripping did reduce most galaxy luminosities. If a large central galaxy was present, it would accrete this material, thereby increasing the luminosity of its halo.

Merritt (1983, 1984) arrived at conclusions that are in striking contrast to the preceding results, based on a statistical description of the galaxy orbital energy and mass distribution and a Fokker-Planck treatment of the evolution of these distributions due to interactions. He found that no significant tidal stripping occurs due to galaxy-galaxy interactions after the period of cluster formation and that galaxy mergers do not occur at a significant rate in rich, virialized clusters. This is a result of strong truncation of galaxies by the mean tidal field of the cluster, which results in galaxies having smaller initial size and thus smaller interaction cross sections. Merritt (1988) thus concludes that, contrary to the "strong" cannibalism hypothesis (i.e, that cD galaxies are produced by merging), the morphology of the FRG is fixed during early stages of cluster evolution and subsequently only "weak" cannibalism occurs, whereby the FRG undergoes only a modest increase in luminosity of a few L^*. Merritt suggests that merging may be far more important in the poor, low-velocity-dispersion groups that merged to form present-day, rich clusters.

Current theories of structure development in a high-density universe ($\Omega \sim 1$) predict that rich clusters grow continuously by accretion of poorer clumps, rather than form "whole" during a single, early collapse and virialization phase. Thus, dynamical processes operating within poor clusters may continuously affect the evolution of rich clusters as these smaller subunits are accreted. The present study has focused on the evolution of poor clusters of 50 galaxies as discussed in § 1.4.

1.3. *Observational Background*

As reviewed by Lugger (1991), a number of observational investigations over the past decade have sought to test the galactic cannibalism theory by investigating the properties of cluster galaxies and their spatial and kinematic distributions within the cluster. Overall, these studies suggest that the effects of dynamical evolution on rich clusters are more subtle than originally predicted. FRGs that are classified as D and cD appear to represent the bright end of the distribution of "normal" FRGs in observed properties, rather than a separate population. This suggests that the same mechanisms—which may include galaxy merger—play a role in the formation and evolution of all FRGs. The improved evolutionary simulations that we report here will allow more specific tests for the effects of dynamical evolution in clusters.

The observation of multiply nucleated FRGs is frequently cited in support of the galactic cannibalism picture. The FRG in as many as 50% of all clusters contains one or more secondary nuclei, i.e., smaller galaxies within 20 kpc (using Hubble constant $H_0 = 50$ km s^{-1} Mpc^{-1}) projected distance of the nucleus of the dominant galaxy (Hoessel & Schneider 1985). The cannibalism picture predicts that the FRG will have satellite galaxies on decaying circular orbits with circular velocities of less than 300 km s^{-1} and orbital radii of 10–20 kpc. Cowie & Hu (1986) analyzed the observed distribution of velocities of 75 secondary nuclei relative to the FRG; they find that the distribution of these velocities is best fit with a two-component Gaussian model in which 60% of the satellites belong to a low-velocity-dispersion population bound to the FRG and the remainder belong to the normal core population of the cluster which has a higher velocity dispersion. Bothun & Schombert (1990) obtained a complete sample of velocities from the inner 600 kpc of eight clusters; the combined sample was also well fit by a double Gaussian, but with 20% of the galaxies belonging to the bound population. They argue that these galaxies are moving on highly elongated orbits, since this population would quickly disappear due to dynamical friction if the galaxies were on circular orbits; a similar conclusion was reached by Tonry (1985).

The practice of combining data from clusters with different velocity dispersions has been criticized by Gebhardt & Beers (1991), who point out that the sum of a number of Gaussians with different dispersions leads to a distribution which is not a Gaussian, but rather is more strongly peaked at low velocities. This is not a difficulty in the approach taken by Lauer (1988), who used photometric data for the central regions of 16 clusters to assess evidence for physical interaction between the nuclei. Half of the sample showed photometric distortions indicative of interaction. These distortions were seen in systems with both large velocity differences between the nuclei (>1000 km s^{-1}) and small differences (<300 km s^{-1}); only the latter group are consistent with cannibalism. Lauer estimates that the central galaxy increases in luminosity by $2L^*$ in 5 Gyr; since a typical cD galaxy has a total luminosity of $\simeq 10L^*$, this is consistent only with weak cannibalism. Similarly, Merrifield & Kent (1991) compute a rate of growth of the central galaxy of only $1L^*$ per 10 Gyr, and find no evidence for a bound population when comparing the velocity distribution of cluster members projected within 20 kpc of the central dominant galaxy with the distribution for control samples drawn from further out in the clusters.

One recent observational development is the discovery of a number of cD galaxies with peculiar velocities larger than the velocity dispersion of the cluster in which they are found (e.g., Zabludoff et al. 1993). To investigate whether this is possible in the strong cannibalism picture, Malumuth (1992) improved on

the scheme of Malumuth and Richstone (1984) by following galaxy orbits in an N-body fashion to identify collision and merging events; only 19% of models with half of the mass in the background formed a cD, a considerably lower rate than seen in the earlier study. Furthermore, the peculiar velocities of the cD galaxies, which decay over time, were too low at the end of the simulations to be consistent with the observed distribution, implying that cD galaxies and/or the clusters in which they are found were formed relatively recently.

1.4. *The Present Study*

As described above, previous computer simulations of the evolution of clusters have tended to rely on approximate cross sections for galaxy interactions, as well as on analytic approximations for the rate of interaction between galaxies and the dark matter background. A full N-body simulation, with each galaxy represented as an extended structure, is clearly preferable since galaxies are continuously undergoing interactions. The rapid increase in available computational power in the past few years has made it possible to introduce this level of physical realism. Most of the simulations reported here were performed on a dedicated HP 9000/735 workstation which runs the N-body tree code at about 20 Mflops. While the simulation of a rich cluster with thousands of resolved galaxies is beyond the capacity of even a high-performance workstation, simulation of a poor cluster, such as those in the MKW (Morgan, Kayser, & White 1975) and AWM (Albert, White, & Morgan 1977) catalogs, is practical. In Bode, Cohn, & Lugger (1993), we reported simulations of groups of 5–10 galaxies; in this paper we present models containing 50 galaxies. The initial conditions and numerical method are discussed in § 2; the results are presented in § 3 and discussed in § 4.

2. THE CLUSTER MODELS

A total of 10 N-body cluster models containing 50 galaxies were evolved. All of the particles in a model have identical mass; a total of $N = 40{,}000$ bodies are used. The models are fully self-consistent in that each galaxy is represented as an extended structure containing many paticles and the gravitational potential arises from the particles alone. Mass is apportioned between the galaxies and a smoothly distributed common group halo, or intracluster background (ICB). The percentage of mass initially in the ICB, β, is chosen to be 50, 75, or 90. Since the total amount of mass in the cluster is the same for all models, increasing β has the effect of removing mass from the galaxies and distributing it throughout the cluster. A snapshot of the initial state of a typical model is show in Figure 1.

To generate the initial conditions, the positions and velocities of the ICB particles and the centers of mass of the galaxies were found by randomly sampling a King distribution with $W_0 = 6$; such a model provides a good fit to existing poor clusters (Malumuth & Kriss 1986). At this stage, each galaxy was temporarily represented by a single particle with a mass equal to the galaxy's total mass and a softening length equal to its half-mass radius. Virial units were used, with gravitational constant $G = 1$ and total mass $M = 1$. The cluster was scaled to have binding energy (ignoring the internal energy of the galaxies) $E = \frac{1}{2}$ and to satisfy the virial theorem, $K - \frac{1}{2}\sum \boldsymbol{F}_i \cdot \boldsymbol{r}_i = 0$, where K is the total kinetic energy and $\boldsymbol{F}_i$ and $\boldsymbol{r}_i$ are the force on and position of the ith particle. With this scaling, the size of the system $R = GM^2/2E$ and initial velocity dispersion $V = (2E/M)^{1/2}$ are both unity; thus the crossing time $T = R/V$ is also unity. Finally, the particles representing galaxies were then replaced with the appropriate galaxy models.

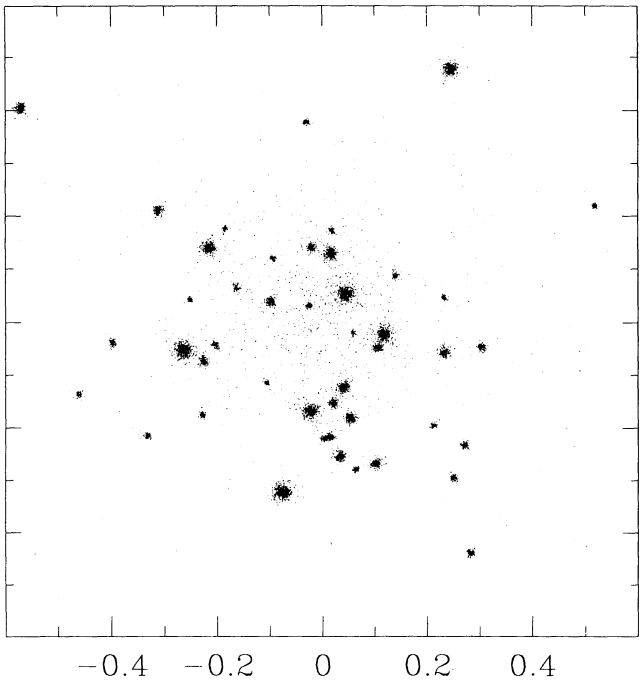

FIG. 1.—Initial configuration of a typical model. Half of the mass is in galaxies and half is in a smooth background ($\beta = 50$). The core radius of the cluster is 0.17 code units, or 250 kpc when scaled to a unit of length of 1.5 Mpc. The total number of particles is $N = 40{,}000$.

As in Bode et al. (1993), the galaxy models consist of a "luminous" core and a "dark" halo. The positions and velocities of the particles are found by sampling a King distribution, the most bound particles are identified as belonging to the core, with the remainder identified as halo particles. The core particles are given a softening length one-fifth of that of the halo particles; the latter are identical to the particles in the ICB. The difference in softening lengths allows the "luminous" matter in the cores to have structure on smaller scales than the more smoothly distributed "dark" halo particles. For a given β, all the galaxies have the same luminous mass fraction. Since the overall luminous mass fraction is 10% for all β, the initial luminous mass fraction for a given galaxy is $0.1/(1 - \beta/100)$. Thus for $\beta = 90$ the galaxies have no dark halos, while for $\beta = 50$ there are 4 times as many dark particles as luminous particles in each galaxy.

The masses of the galaxies follow a Schechter distribution function

$$\phi(n) = \frac{dN_g}{dn} = \frac{\phi^*}{N^*}\left(\frac{n}{N^*}\right)^{-1.25} e^{-n/N^*}, \quad (1)$$

where N_g is the number of galaxies, n is the number of particles in a galaxy, and N^* is the number of particles in a galaxy of characteristic mass M^*; ϕ^* is a normalization parameter. The range of masses was constrained by the desire for resolution at the low end and the need to limit the total size of the model at the high end. The smallest galaxy was chosen to have 25 luminous particles, and an N^* galaxy to contain 175. The masses

were not drawn randomly from the distribution, but rather by integration of equation (1), with the number of galaxies, N^*, and the minimum and maximum sizes fixed beforehand. Since a significant fraction of the cluster mass is in a small number of large galaxies, it was considered best to fix their initial masses in this way. With the largest galaxy containing 347 luminous particles, the total number of luminous particles is 4004. There is a factor of 14 in mass between the largest and smallest galaxies, which if M/L is independent of galaxy size would be a difference of not quite 3 mag. Comparison of the chosen range to the total distribution shows that 76% of the total mass is included.

The binding energy per unit mass for an N^* galaxy is set equal to that of the cluster as a whole, or $E_g/M_g = 0.5$ for $M_g = M^*$. The other galaxies are scaled such that $E_g/M_g \propto M_g^{1/2}$; this means the galaxy velocity dispersion and radius scale roughly as $v_g \sim M_g^{1/4}$ and $r_g \sim M_g^{1/2}$, respectively. Thus the larger galaxies are more tightly bound and have internal velocity dispersion approaching that of the cluster as a whole. The King central potential parameter W_0 is chosen so that when the core particles of an N^* galaxy are considered in isolation their binding energy per unit mass also equals $\frac{1}{2}$. This ensures that the cores have radial luminous mass profiles that are independent of the halo mass (i.e., independent of β). Individual galaxy models were evolved for three time units before being placed in the group model. Test galaxies were also evolved in isolation for the same length of time as they would be in the cluster models in order to check for stability. There was little change seen in the galaxy parameters, the largest being an increase in the radius containing 90% of the luminous mass in the smallest galaxy of about 5% over eight crossing times.

The distribution of particles between the galaxies and ICB is summarized in Table 1. In code units, the softening length of the core particles is $\epsilon_l = 0.002$ and that of the halo particles is $\epsilon_d = 0.01$ for the $\beta = 50$ models, and $\epsilon_l = 0.0018$ and $\epsilon_d = 0.009$ for the other models. This high degree of softening is necessary to prevent two-body relaxation, given the relatively small number of particles in each galaxy. The core radius of the clusters is ≈ 0.17. Scaling the units of mass and length to $10^{14}\ M_\odot$ and $R = 1.5$ Mpc makes the unit of time $T = R/V = (R^3/GM)^{1/2} = 2.7$ Gyr; the unit of velocity is then $V = 535$ km s^{-1}. We use $H_0 = 50$ km s^{-1} Mpc^{-1} throughout this paper. Using this scaling, the core radius is $R_C \simeq 250$ kpc and the three dimensional velocity dispersion is $V \simeq 535$ km s^{-1}. The sizes and dispersions taken on by representative galaxies under this choice of scaling is shown in Table 2. The maximum size of a galaxy in a cluster will be limited by the tidal force of the cluster; this tidal radius is given by

$$\frac{R_T}{R_C} \simeq \frac{1V_g}{2V} \qquad (2)$$

TABLE 1

PARAMETERS FOR 50-GALAXY MODELS

ICB Mass Fraction β	Number in Galaxies	Number in ICB	W_0 of Galaxies	Runs
50.....	20010	19990	8.25	5
75.....	10006	29994	6.25	2
90.....	4004	35996	3.00	3

TABLE 2

GALAXY PARAMETERS

		Total				Core Only			
β		n	$r_{1/2}$ (kpc)	$r_{90\%}$	$\langle v \rangle$ (km s^{-1})	n	$r_{1/2}$ (kpc)	$r_{90\%}$	$\langle v \rangle$ (km s^{-1})
50.....	M_{max}	1734	22	65	550	347	4	9	560
	M^*	875	16	48	410	175	3	5	420
	M_{min}	125	5	11	124	25	1	2	100
75.....	M_{max}	867	11	28	510	347	4	9	550
	M^*	438	8	19	390	175	3	6	430
	M_{min}	63	2	6	96	25	1	2	95
90.....	M_{max}	347	5	10	520	347	5	10	520
	M^*	175	3	7	410	175	3	7	410
	M_{min}	25	1	2	110	25	1	2	110

(Merritt 1984). Since the large galaxies have a dispersion $V_g \approx V$, the limiting tidal radius is $R_T \approx 125$ kpc, or about twice the largest initial galaxy size.

The evolution was carried out using a modified version of Hernquist's "tree" code (Hernquist 1987, 1990). The tree code is well suited for simulating inhomogeneous systems with little spatial symmetry such as galaxy clusters. The changes made to Hernquist's code are discussed in Bode et al. (1993); one of these is the use of a different cell opening criterion: a cell is used if

$$d > 2\epsilon_d + \sqrt{3}\, s\,, \qquad (3)$$

where d is the distance from the particle to the center of mass of the cell and s is the size of the cell. This is basically the standard Barnes-Hut criterion with tolerance parameter $\theta = 1/3^{1/2} \simeq 0.58$, but using $d - 2\epsilon_d$ rather than d. Spline-kernel softening is used.

For the $\beta = 50$ models, a single step took roughly 35 CPU s on a Cray Y-MP (running at 118 Mflops), 210 s on a HP 9000/735, and 318 s on an IBM RS-6000/560; higher β models require fewer interaction terms and thus less CPU time. The average number of interaction terms used for calculating the force on a particle is initially about 1200 (for $\beta = 50$), increasing slowly throughout the evolution to over 1400; the other models used 15%–20% fewer terms. The tree code does not exactly conserve energy or momentum (Hernquist 1987); with time step size $\Delta t = 5 \times 10^{-4}$, energy was conserved to better than 0.1% over an entire evolution, and the center of mass (originally at the origin) moved a distance of order 10^{-3} code units. Four of the $\beta = 50$ models were run with a time step of $\Delta t = 10^{-3}$, which conserved energy to $\lesssim 1\%$ and conserved momentum to a similar degree as the models with smaller Δt; the other $\beta = 50$ model, run with $\Delta t = 5 \times 10^{-4}$, did not behave in any markedly different fashion from the others.

3. THE EVOLUTIONS

The five $\beta = 50$ models all followed a similar pattern of behavior. Snapshots from typical runs with different β are shown in Figures 2–6. The galaxy number density near the center of the cluster tends to rise during the first two crossing times. At this point the merging rate increases; merging tends to occur near the center of the cluster and involves the more massive galaxies. Within four time units ($\approx 10^{10}$ yr) this merging has resulted in the formation of a dominant, centrally located galaxy. Almost all of the subsequent merging involves this dominant galaxy accreting the others. By the end of the

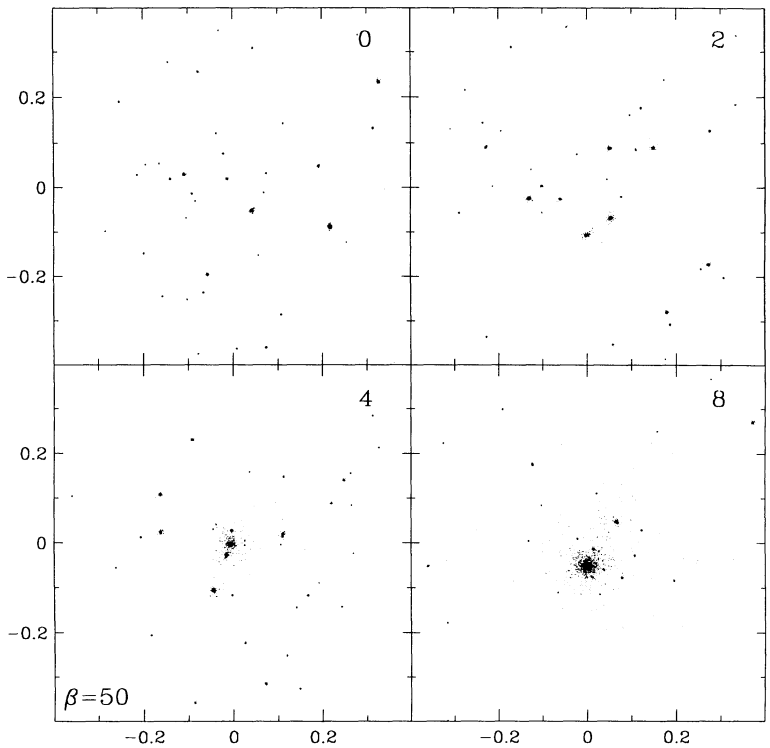

FIG. 2.—Snapshots showing the inner region of a model. The time is given in the upper right-hand corner; four time units equal roughly 10^{10} years. Only the luminous cores of the galaxies are shown; the dark galaxy halos and the cluster background are not. Half the mass is in the cluster background.

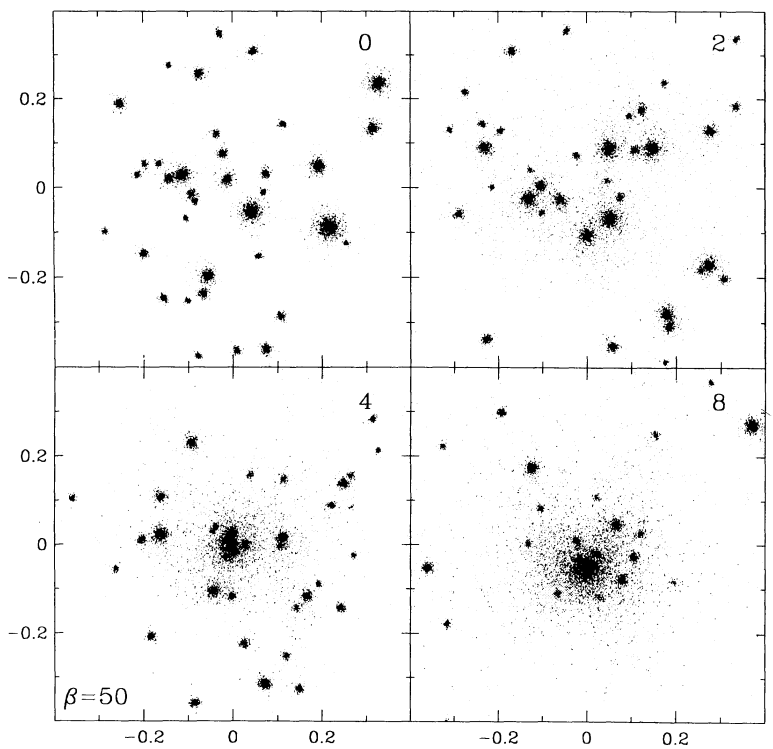

FIG. 3.—Same as Fig. 2, but showing all the mass initially attached to the galaxies, including the dark matter halos. The particles initially in the ICB are not shown.

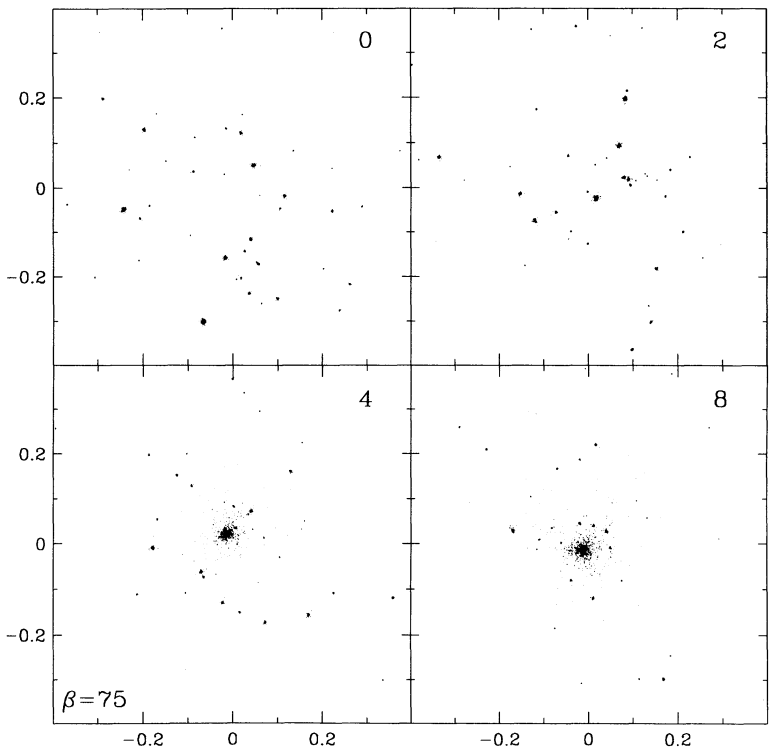

FIG. 4.—Same as Fig. 2, but for $\beta = 75$

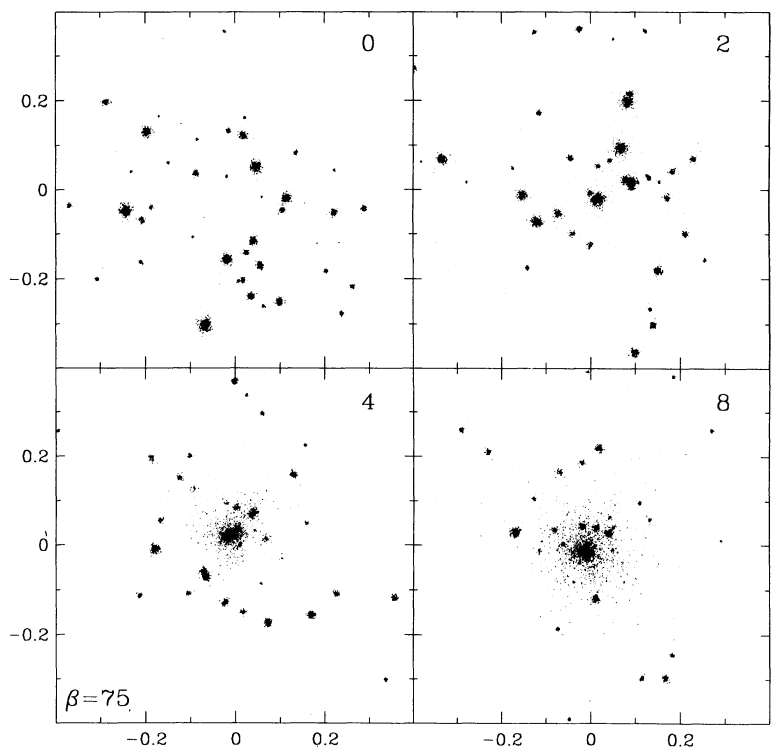

FIG. 5.—Same as Fig. 3, but for $\beta = 75$

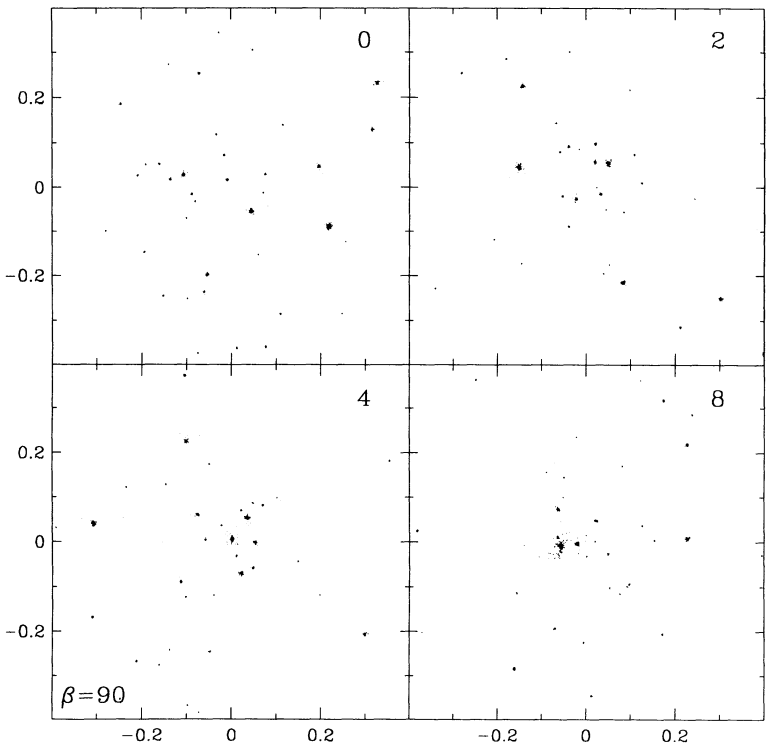

FIG. 6.—Same as Fig. 2, but for $\beta = 90$. The galaxies do not have dark halos.

evolution this galaxy contains a third or more of the luminous matter in the cluster.

The positions and velocities of the particles were saved every 0.025 time units for analysis. Only the luminous particles were used in calculating galaxy positions and velocities. This was done using the "friends-of-friends" algorithm, in which particles are linked if they are closer together than some separation parameter, d. An additional requirement was made that the velocity difference between the two particles be less than the initial velocity dispersion of all the luminous particles. Particles in each galaxy are thus linked together either directly or indirectly. Groups of particles with fewer members than an adopted cutoff (chosen to be $\simeq \frac{2}{3}$ of the size of the initially smallest galaxy) were not considered as galaxies. The separation parameter was chosen such that 75% of the particles were included in some galaxy, leading to a distance $d \sim 0.003$. This method was able to distinguish galaxies undergoing close encounters without making spurious identifications of small subgroups. One drawback of this technique is that, while all of the particles in the smaller galaxies are located, particles that are located in the outermost, low-density regions of the largest galaxies are not counted. Thus for a galaxy with an extended luminous envelope, only the particles in the core will be identified as belonging to it.

3.1. *Merging*

The merging histories of the five $\beta = 50$ models are given in Figure 7, which shows the number of galaxies as a function of time for each run. It can be seen that there is little variance in the final state between runs; there are between 14 and 17 mergers by $t = 8$. Although the average merging rate is quite similar between models, within a model the rate can vary significantly over time—there can be several mergers within a crossing time or a period of over a crossing time with no merging. There are very few mergers within the first crossing time. In two models a merger occurs almost immediately, involving galaxies that are initially bound to each other; there

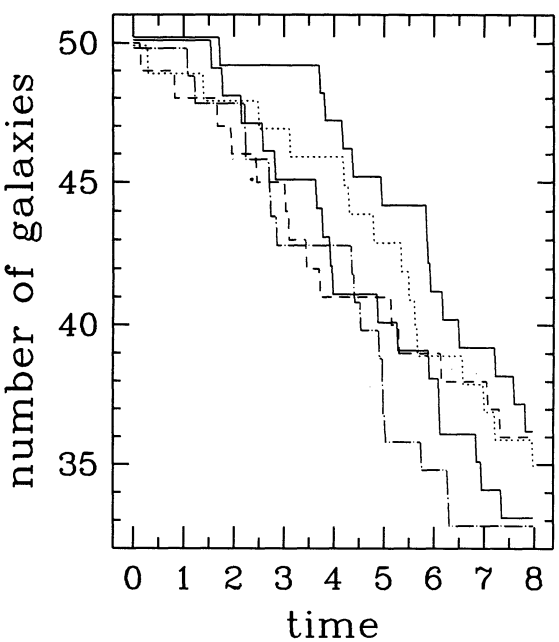

FIG. 7.—Number of galaxies as a function of time for the five models with 50% of the mass in galaxies ($\beta = 50$); the number decreases as a result of mergers. The five models differ only in the randomly generated initial galaxy positions and velocities, which were drawn from a King-model distribution. Small vertical shifts were applied to the curves for clarity

is only one other merger before $t = 1$. Dynamical friction causes a general trend toward rising number density in the inner 100 kpc, as galaxies (particularly the most massive) lose orbital energy; the central density generally peaks between $t = 4$ and $t = 5$ and falls off afterward due to merging.

The mean number of surviving galaxies in these five runs is shown as a solid line in Figure 8. After the initial period of slow merging, the merging rate is roughly two mergers per unit time, slowly increasing to three per time unit by $t = 5$, and dropping toward the end of the simulation. Figure 8 also shows the merging histories for models with higher β. The two $\beta = 75$ models have quite similar merging histories. There is a longer interval at the outset of the simulation where no merging occurs, followed by a burst of rapid merging at $t \approx 3$ such that by $t \approx 4$ the total number of mergers is as great as is seen in some of the $\beta = 50$ models. The merging rate then drops to a lower value of one to two mergers per crossing time.

The three $\beta = 90$ models are also shown. Here there is a wider range of behavior than seen previously. The trend of delayed onset of merging with higher β is continued. One model behaves in a fashion similar to the $\beta = 75$ models, while the others show no significant merging until near the end of the simulation (after 13 Gyr). When the merging finally does begin, it is at a rate similar to the other models.

As β is increased, galaxy halo sizes are decreased, thus reducing both the initial mass and geometric cross section. A decrease in the cross section of the galaxies allows a greater degree of orbital decay to take place before close encounters result in merging. Also, a decrease in the initial mass of the galaxies results in longer dynamical friction timescales. Thus the trend of increasing delay in the onset of merging is consistent with expectation.

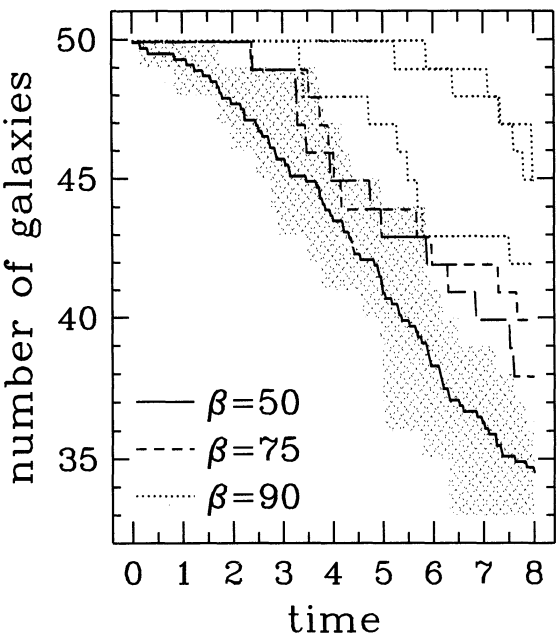

FIG. 8.—Number of galaxies as a function of time. The curves are labeled by the percentage of mass in the cluster background, β. The hatched area shows the range of values seen in the five $\beta = 50$ models, while the solid line is the mean value. The dashed lines are two $\beta = 75$ models; the dotted lines are three $\beta = 90$ models.

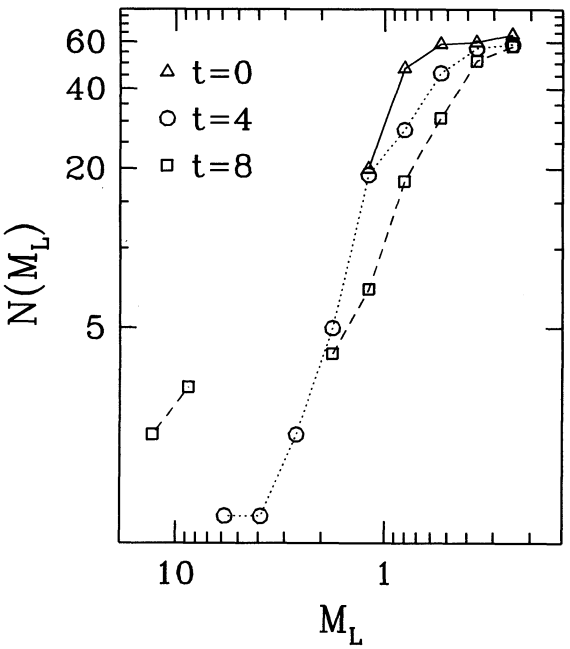

FIG. 9.—Luminous mass function obtained by summing together the five $\beta = 50$ models, in units of M_L^*. Times are given in the upper left-hand corner. Only the "luminous" particles in the cores of the galaxies are used to determine the mass.

3.2. *The Mass Function*

The largest galaxies preferentially engage in merging—few mergers involve the least massive galaxies. This can be seen in Figure 9, where the differential mass function is shown. The mass is taken to be the number of luminous particles in the galaxy found by the friends-of-friends algorithm normalized to the initial number found in an N^* galaxy. At each epoch all five $\beta = 50$ models are binned together. It can be seen that the number of smaller galaxies changes little over time; instead, the larger galaxies merge, initially extending a roughly power-law tail to higher masses, which eventually turns into a steep cutoff increasingly separated from a single large merger remnant. This latter phase is the type of evolution predicted by Hausman & Ostriker (1978).

By $t = 4$ (10 Gyr), the largest galaxy has accreted 0.5–5.0L^*. The difference in magnitude between the two brightest galaxies changes little before $t = 2$; after that point the mean difference increases roughly linearly in time by 2 mag over the next six time units. The increasing difference is due in part to the growth of the FRG; it also rises when the two largest galaxies merge, so that a smaller galaxy becomes the second-ranked one. By $t = 8$ the luminosity function is no longer continuous; there is a large gap between the FRG and the rest of the distribution. The five largest galaxies shown in the $t = 8$ curve are the FRG from each of the five models; they are a factor of ≈ 6 more massive than the second-ranked galaxies. The scatter between models in the size of the FRG is surprisingly small, the largest being less than a third larger than the smallest.

The luminosity function of the $\beta = 75$ models does not evolve as quickly, and the final masses of the FRG are smaller. For the $\beta = 90$ models much less change is apparent. In the most extreme case, the FRG gains less than $3L^*$ over the entire run; the near lack of merging in the other models precluded any growth.

3.3. *Mass Segregation*

If mass segregation is taking place, different mass groups would have different harmonic mean separations, given by

$$r_h = 0.5N_g(N_g - 1)\left(\sum_i \sum_{j<i} \frac{m_i m_j}{r_{ij}}\right)^{-1}, \quad (4)$$

where r_{ij} is the distance between two galaxies i and j, and N_g is the number of galaxies summed over. Galaxies were divided into two groups depending on whether they were found to contain more or fewer luminous particles than $N^*/2$, so that at the initial step the high-mass group contained 15 galaxies. The initial value of r_h is 0.5 code units (750 kpc) for both groups. For the lower mass group the mean value of r_h changes little with time, but in the higher mass group it falls quickly, so that by $t = 1$ the median $r_h \approx 0.35$ code units. After $t = 5$ it begins to rise again as merging reduces the number of close pairs of galaxies contained in the group. Yepes, Dominguez-Tenreiro, & Pozo-Sanz (1991) found variations of 20% or less between the characteristic lengths between various luminosity groups in three rich clusters. A considerably larger variation is often found in our simulations; however, there is scatter in the r_h seen for each mass group with a standard deviation of roughly 10% With as few as 50 galaxies there is enough statistical fluctuation that mass segregation will not always be evident.

The velocity dispersions for both mass groups were also calculated. While the dispersion of the high-mass group changes little in time, that of the low-mass groups rises 10% by $t = 3$, leveling off afterward. Since the dispersion is initially unity in code units, this is an increase of 50 km s^{-1} using our adopted scaling. Since the dispersion in a given model varies by 50–100 km s^{-1} due to statistical fluctuations, it is questionable whether such a difference could be observed. Biviano et al. (1992) found significant segregation in velocity only when considering the three brightest galaxies.

3.4. *FRG Position and Velocity*

Initially the largest galaxy is simply the extreme end of the mass distribution and may be anywhere in the cluster. As larger galaxies move to the core and are involved in merging, one of these merger remnants will become the largest galaxy. Thus, once merging begins, this FRG will likely be found near the center of the cluster. Dynamical friction will cause the FRG's orbit to further decay. This can be seen in Figure 10, which shows the distance of the FRG from the center of mass of the cluster as a function of time. The curves begin at $t = 3.5$, when more than two models contain a FRG with mass greater than $2M_i$, where M_i is the mass of the largest galaxy at $t = 0$. The mean distance, when the large galaxies are formed at $t \gtrsim 3$, is over 0.1 code units or 150 kpc; it then drops over the next time unit to less than 0.05 code units. The maximum distance seen in the five models begins at twice the mean and declines to near the mean at less than 100 kpc. Beers & Geller (1983) find the mean distance of D and cD galaxies from local density peaks to be 90 $\pm$ 25 kpc. This is in good agreement with the models, excepting the earliest and latest times.

The peculiar velocities of the FRG show little secular change after $t \approx 4$. Once the largest galaxy has grown to $2M_i$, its velocity has generally decreased to below 0.7 code units (270 km s^{-1}). Typical line-of-sight peculiar velocities are $\approx$0.2 code units (170 km s^{-1}). There can be occasional peaks in these velocities (over 400 km s^{-1}) during the first couple of crossing times after formation, as encounters and mergers perturb the FRG's orbit. As the FRG becomes larger and the perturbing galaxies become smaller on average, these perturbations have less effect, so that the maximum FRG velocities decrease. However, the median value remains near 170 km s^{-1}.

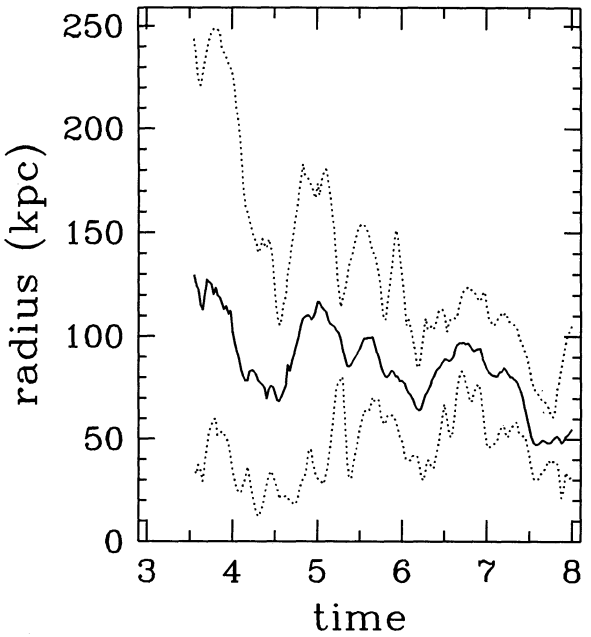

FIG. 10.—Distance of the most massive galaxy from the center of the cluster as a function of time, averaged over the $\beta = 50$ models. The curves begin when the largest galaxy is over twice as massive as the initial largest galaxy mass in more than two of the models. The dotted curves show the minimum and maximum values, while the solid line is the mean value. The core radius of the clusters is 250 kpc.

A measure of the peculiar velocity of D/cD galaxies in clusters advocated by Gebhardt & Beers (1991) is the normalized offset relative to scale, or Z-score, given by

$$Z = \frac{V_{cD} - C_{BI}}{S_{BI}}, \quad (5)$$

where V_{cD} is the velocity of the FRG, C_{BI} is the biweight estimator of the central location in velocity space, and S_{BI} is the biweight estimator on scale in velocity. The Z-score is thus analogous to the deviation from the mean normalized by the dispersion. The details of the estimators C_{BI} and S_{BI} are given in Beers, Flynn, & Gebhardt (1990). Girardi et al. (1993) have shown, in a study of 79 clusters, that when more than 20 galaxy redshifts are available S_{BI} behaves similarly to the dispersion. When fewer velocities are available, the robust biweight estimators are more efficient than the classical mean and standard deviation.

For three orthogonal projections, the Z-score of the FRG was calculated every 0.025 time units using the ROSTAT Fortran routines written and tested by Beers et al. (1990). The resulting distribution of Z-scores is shown in Figure 11. Initially the FRG velocities will be similar to those of other cluster galaxies, so $Z > 1$ is common. Due to dynamical friction and merging, the FRG is found in the core with a lower velocity at later times. If we consider only those times after the point at which the FRG has grown to a mass $M_{FRG} > 2M_i$, then $Z > 1$ is quite rare. The FRG reaches this mass cutoff at $t \gtrsim 3.5$; there is little evolution in the kinematic status of the FRG once this point is reached. There are slightly fewer high Z-scores ($Z \gtrsim 1$)

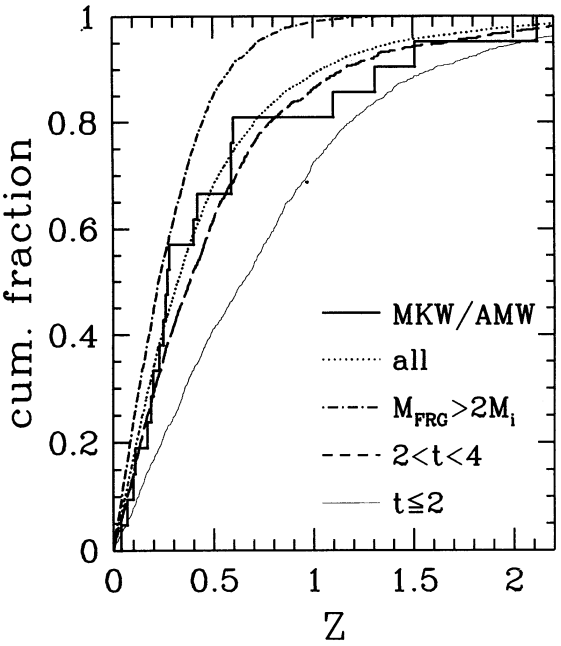

FIG. 11.—Cumulative distribution of Z-scores. *Thick line*, data for MKW/AWM clusters; *dotted line*, all times from the models; *dot-dashed line*, when the FRG is over twice as massive as the largest galaxy at $t = 0$; *dashed line*, times $2 < t < 4$; and *thin line*, earliest times.

but the rest of the distribution remains unchanged. The higher β models show similar Z-score distributions when the mass cutoff is used; in all cases the maximum Z-score ever seen is near unity.

Also shown, as a thick solid curve, in Figure 11 are data from Beers et al. (1994), who have compiled enough redshifts in 21 MKW and AWM clusters to compute Z-scores for the D/cD galaxy. The resulting Z-scores tend to be higher than those in our models at later times. While the median Z-score in the models (when $M_{\rm FRG} > 2M_i$) is 0.22, for the poor clusters it is 0.27. The biggest difference is at higher Z; while four of the 21 clusters have $Z > 1$, this almost never occurs in the models. The distributions differ at the 99% significance level (using a K-S test). If we eliminate the four Z-scores which are greater than unity from the data, then the fit is improved, but there is no compelling reason to consider these four clusters as being different from the rest. Such large Z-scores may not be limited to smaller groups. In the sample of richness class $R \geq 1$ Abell clusters of Zabludoff et al. (1993) three out of 25 cD's have peculiar motions greater that the cluster velocity dispersion; a third of the cD's have peculiar velocities larger than half the cluster dispersion. On the other hand, Bird (1994) found $Z < 0.4$ in 25 clusters with at least 50 measured redshifts. These higher peculiar velocities could be the result of some external influence of larger structures or ongoing merging with other clusters; if this were the case, then one should be able to find photometric and kinematic evidence of such interactions in the neighborhoods of these clusters.

Alternatively, the clusters may have a variety of ages, with those showing high Z-scores being among the younger ones. Interestingly, if we take the Z-scores of the $\beta = 50$ models for all times, regardless of the mass of the FRG, the result is a much better fit to the data, as is shown as a dotted line; these two curves appear to arise from the same distribution. A significant spread of ages would require that the D/cD formation process take place by or soon after the time that the cluster is virialized. In this case, the high-Z clusters may again be distinguishable from the others since they would be dynamically younger. This spread of ages need not include the entire span of our simulations; if we consider the Z-scores in a smaller time range, $2 < t < 4$ (disregarding the mass of the FRG), the cumulative distribution is quite similar to the curve for all t. This suggests a median age of ~ 8 Gyr using our adopted scaling.

An associated question is how large a Z-score should be before it can be considered significant. The 90% confidence intervals determined by 1000 bootstrap resamples was found to work reasonably well. In our models this means FRGs with true peculiar velocities less than approximately 60 km s^{-1} (relative to the center of mass) do not usually have statistically significant Z-scores, while those with true velocities greater than 100 km s^{-1} do have statistically significant Z-scores 95% of the time.

3.5. *The Density Distribution*

To measure the density profile, the galaxy positions found in the manner described above were divided into radial bins centered on the center of mass. The profiles from all models with the same β were averaged together to improve the statistical weight. The resulting density profiles were also averaged over a number of times. The FRG was always excluded from the bin counts, given the special dynamical status it achieves.

The solid line in Figure 12 shows the resulting three-dimensional density distribution at the beginning of the simulations (averaging over all times before $t = 1$). The radial bin size is 50 kpc and the inner 500 kpc is shown; the change in slope at the core radius $R_C = 250$ kpc is clearly visible (the results have been scaled to physical units as described in § 2). The squares show the number density averaged over times

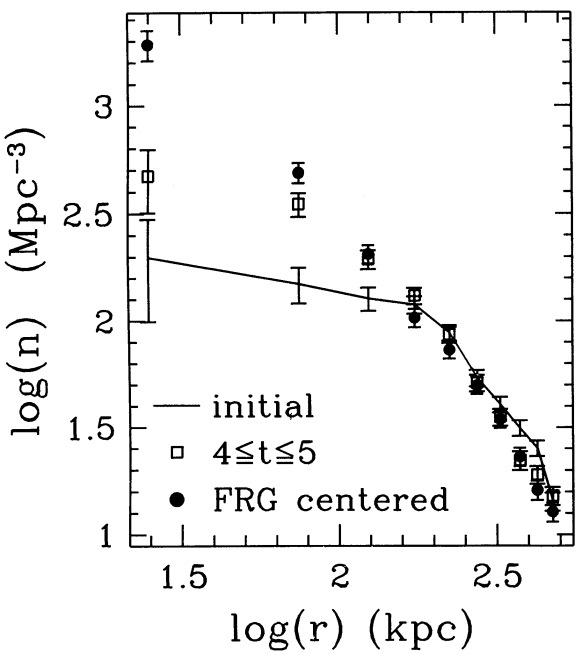

FIG. 12.—Three-dimensional galaxy number density found by averaging together the five $\beta = 50$ models. *Solid line*, initial state ($t < 1$); *open squares*, state at $4 \leq t \leq 5$; and *filled circles*, same as the open squares but calculated using the FRG as the cluster center. Error bars show 2σ, where σ is taken to be the square root of the number of counts in each bin.

$4 \leq t \leq 5$; little change occurs outside the core aside from a slight drop in density. The most evident change occurs in the inner 200 kpc: it appears much as if the core radius has become smaller while keeping the King-model profile. Remember, however, that the FRG is excluded—by $t = 4$ it has fallen to the center of the cluster, and counting it would increase the number density in the inner 100 kpc. If the FRG is included, then the core is even less apparent.

In observational studies, the FRG is often taken to be the putative cluster center. The filled circles in Figure 12 correspond to the squares, except that all radii are measured relative to the FRG; that is, the FRG is taken to be the "center" of the cluster rather than the actual center of mass. The resulting profile approaches a power law, very nearly an isothermal sphere. Excluding the innermost bin, the best linear least-squares fit in the log-log plane has a slope of -2. If the presence of the FRG is included in the innermost bin then it also is well fitted by the r^{-2} profile. The profiles achieved by $t = 5$ change little over the rest of the simulation.

Of course, clusters are actually seen in projection; the density profile evolution seen in three dimensions is also visible in two. Figure 13 shows the suface number density of the models in projection. Three orthogonal projections were taken from each model and averaged together; after that the calculations were identical to those in computing the three-dimensional densities. The radial bin size is 25 kpc. As before, the existence of a core is obvious at $t = 0$, and by $t = 4$ it has been partially erased, with the density falling slightly outside the core and rising inside. If we take the FRG to be the center of the cluster, then the core is largely erased; when the FRG is included in the central bin, the surface density profile is nearly an r^{-1} power law. Beers & Tonry (1986), in a study of 48 Abell clusters, conclude that rich clusters do have an r^{-1} profile when the location of either the D/cD galaxy or the X-ray maximum is used as the cluster center. The main change in the profile after $t = 5$ is that the density in the central bin decreases to $\frac{2}{3}$ of its maximum value (at $t = 4$–5). A similar drop is seen in the space density if we use a bin size of 25 kpc.

Merrifield & Kent (1989) stacked together CCD images of 29 Abel clusters of redshift $z < 0.1$ whose FRG had multiple nuclei; the FRG was used to define the cluster center. The resulting number density profile within a projected radius of 500 kpc is well described by the exponential distribution

$$\mu(R) = \mu_0 e^{-R/R_{\rm exp}} . \quad (6)$$

The value of the e-folding length $R_{\rm exp}$ appears to be independent of cluster richness, the mean best-fit value being 188 kpc. In practice the value of the central density μ_0 depends upon the limiting absolute magnitudes and degrees of completeness of the position data from the cluster images. Assuming a Schechter form for the luminosity function and an apparent magnitude cutoff of 20 in all the clusters, Merrified & Kent (1989) arrive at a mean normalized projected central density of $\mu_0 = 300$ Mpc^{-2} for richness class 0 and 1 clusters, with a dispersion of 100 Mpc^{-2}.

Our projected data set (using the FRG as the center of the cluster) can also be well fit by this form, as is shown in Figure 14. Oddly enough, this is true even at early times when the FRG is not necessarily near the cluster center, although with an e-folding length of more than 500 kpc. After time $t = 4$, when the FRG has assumed a central position, the best-fit e-folding length stays roughly constant at $R_{\rm exp} \approx 155$; the central density μ_0 found from the fit is $\mu_0 \approx 190$ Mpc^{-2} at $t = 4$, declining over time at a rate of 10 Mpc^{-2} per unit time. However, by $t = 5$ the density rises more steeply in the core, such that the density in the inner 50 kpc is higher than the fit by 10%. Using the true center of mass rather than the position of the FRG also results in a profile well fit by an exponential, although the fit improves as the FRG spirals in closer to the center. The central density is lower ($\approx$150 Mpc^{-2}) and

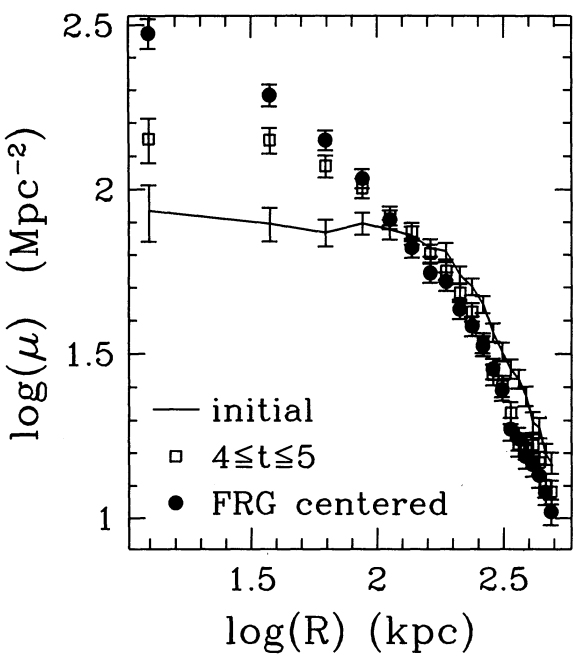

FIG. 13.—Surface galaxy number density. The symbols have the same meaning as in Fig. 12.

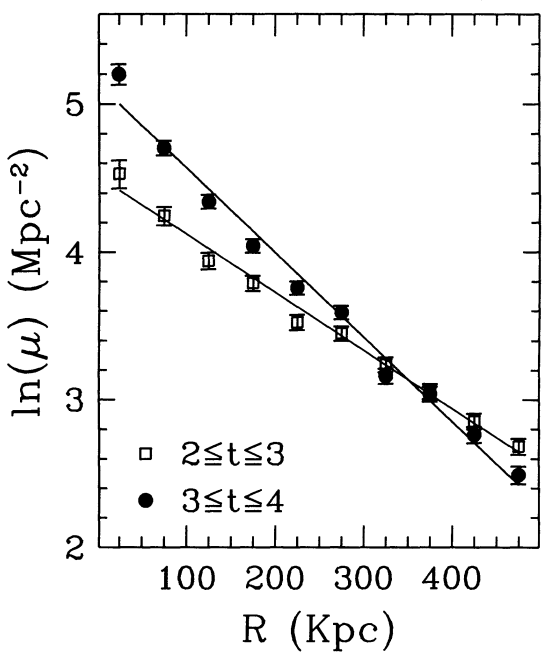

FIG. 14.—Surface galaxy number density, in the ($\ln \mu$, R)-plane. A least-squares linear fit is also shown for each epoch. After $t = 4$, the fit becomes poorer since the central regions rise more steeply.

e-folding lengths are longer (declining from 180 kpc at $t = 4$ to 160 at $t = 8$).

Moving from $\beta = 50$ to $\beta = 75$ makes little qualitative difference in the density evolution The FRG is smaller and tends to remain closer to the true center of mass. Using either the center of mass or the position of the FRG produces very similar results after $t = 4$, with the FRG-centered density being significantly higher in the inner 25 kpc only. The increase in density inside the core is about a third less than in the $\beta = 50$ case and is not as sharply peaked. In the $\beta = 90$ models, using the FRG as the cluster center does not produce a power-law profile; instead, only using the true center of mass results in a strongly peaked profile. However, the evolution is so slow that the density profile continues to steepen throughout the simulation. By $t = 8$ the density profile behaves roughly as r^{-2}, with a central density similar to that in the $\beta = 50$ models when centered on the FRG.

3.6. *Multiple Nuclei*

The density distribution is closely related to the existence of multiple nuclei. Observationally, these are usually defined as a second nucleus within 20 kpc (projected) of the FRG. As before, we examined each of our models every 0.025 time units from three orthogonal projections, looking for galaxies with projected separations from the FRG of less than 20 kpc (0.013 code units).

Figure 15 shows how often such a secondary nucleus occurs as a function of time, the results being expressed as a percentage of the snapshots examined in each time interval. For the first three crossing times, the orbit of the FRG is decaying to a mean radius well inside the core. Since it spends much of its time in the less populated outer regions of the cluster, there are typically few nearby galaxies initially. By $t = 2$–3, the orbit of the FRG lies completely inside the core; the percentage of snapshots with multiple nuclei seen at this point is similar to the frequency of galaxies seen projected within 20 kpc of the true center of mass. This latter number is in the range 10%–15% throughout the simulation. After $t = 3$, multiple nuclei are seen in at least 20% of the snapshots. The frequency rises to over 37% during times $t = 4$–5 and falls off afterward.

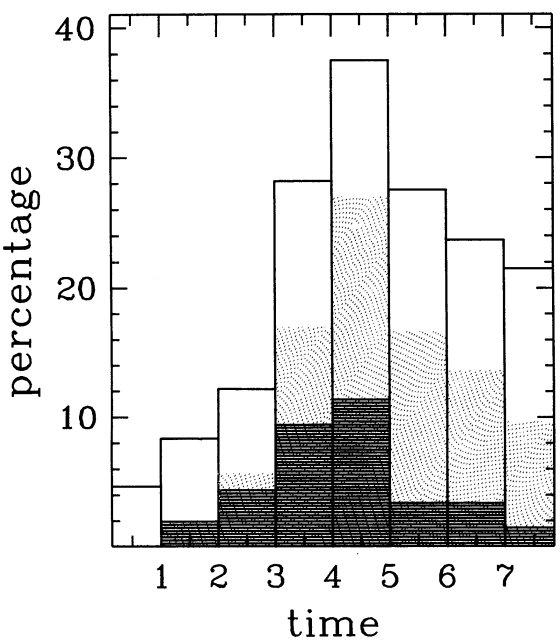

FIG. 15.—Percentage of FRGs with secondary nuclei versus time for $\beta = 50$. The unshaded bars are for a projected two-dimensional separations of less than 20 kpc; the dark-shaded bars show those nuclei that also have true three-dimensional separations of less than 20 kpc. The light-shaded bars show those nuclei that will merge with the FRG in less than 2 Gyr.

In order to clarify the dynamical status of the secondary galaxies, Figure 15 also shows the percentage of models where there is a companion to the FRG with a true three-dimensional separation of 20 kpc or less. Like the projected separations, the number of true companions increases until $t = 5$ and falls off afterward. If a galaxy passes this close to the FRG, it merges with the FRG within a short interval—almost always within the next 2 Gyr. (There are exceptions to this for $t < 3$, when the FRG is not as large and is moving faster.) Once the FRG assumes a position inside the core, it begins merging with the galaxies it encounters until it clears out the portion of phase space containing orbits that intersect with its position.

The merging rate does not change as dramatically as the drop seen in true close encounters might suggest. To explore this further, Figure 15 also shows the percentage of multiple nuclei which are destined to merge with the FRG within the next 2 Gyr. (Note that this means none of the snapshots from the final 2 Gyr can be counted, even though they might have been had the evolution been carried further.) This number also rises to a peak at $t = 4$–5 and falls off, but more slowly. As the FRG consumes the nearby galaxies, the number of multiple nuclei decreases. As the FRG grows in size, a victim galaxy will not survive as close a passage as before. Indeed, at later times, galaxies in some instances are captured from beyond 20 kpc in less than 0.025 time units (less than 10^8 yr).

After $t = 3$, the percentage of snapshots which have multiple nuclei which are not going to merge in the next 2 Gyr is roughly constant at about 11%. These are drawn from the background distribution and are projections of galaxies from the outer parts of the cluster. As those galaxies which do in fact come close to the FRG are consumed, fewer are left, thereby causing a drop in the percentage of secondary nuclei seen. The decrease is not as dramatic as is the drop-off in true three-dimensional encounters. The increasing size of the FRG means that satellite galaxies will spend less time inside a 20 kpc radius from its center than before.

How does the number of multiple nuclei seen compare with expectations? For times $3 < t < 7$ there are an average of 12.5 galaxies inside the core ($r < 250$ kpc). Assuming roughly constant surface density inside the core, such as an analytic King model

$$\mu(R) = \frac{\mu_0}{1 + (R/R_C)^2}\,, \tag{7}$$

leads to the expectation of a galaxy within $r = 20$ kpc 11% of the time. While the models begin with this sort of profile, by $t = 4$ this is no longer the case, as is shown in Figures 12 and 13. A more strongly peaked profile which provides a better fit to the inner regions, such as equation (6) with $R_{exp} = 155$ kpc and $\mu_0 = 190$ Mpc^{-2}, predicts 14 galaxies inside 250 kpc, and a galaxy within $r < 20$ kpc 22% of the time. Since the best-fit μ_0 declines over time to $\mu_0 = 160$, the frequency of multile nuclei is also expected to decline to 19%. Comparing this prediction to Figure 15 shows that for $t > 6$ there is a good agree-

ment with the actual number. However, for $t = 3$–6, and especially for $t = 4$–5, considerably more multiple nuclei are seen than are expected. The spatial distribution of galaxies is even more strongly peaked than the exponential profile, as can be seen from Figure 14. However, the strong central peak is most pronounced at earlier times ($t = 4$–5). This peak is a result of the initial relaxation of the distribution. Once candidate galaxies are used up from this initial burst, the merging rate slows.

The velocities of the secondary nuclei are isotropically distributed. Considering those satellite galaxies which do merge with the FRG at some later point, there is a trend from more radial orbits at earlier times to more circular orbits at later times. For times $3 < t < 5$, 25% of these premergers have orbits with respect to the FRG such that

$$\frac{\boldsymbol{v} \cdot \boldsymbol{r}}{rv} > 0.8 , \tag{8}$$

while for times $6 < t < 8$ only 8% satisfy such a condition. Those galaxies on plunging oribts are quickly consumed by the FRG and are replaced at later times by galaxies circling in as they lose orbital energy through dynamical friction. Merrifield & Kent (1989) argue that galaxies in the process of a direct collision with the FRG could cause multiple nuclei in about 10% of clusters; this level is consistent with the additional number of nuclei seen in our models.

The line-of-sight velocities of the secondary nuclei are generally well fit by a Gaussian with dispersion $\sigma \approx 1$ in code units. Since the three-dimensional dispersion of the cluster as a whole is ≈ 1, this means the secondary nuclei are moving faster than average, as would be expected for galaxies near orbital pericenter. Secondary nuclei may have quite large velocities even though they may be about to merge with the FRG. Those secondary nuclei that are going to merge within the next Gyr have a distribution of velocities similar to that of the other nuclei, such that line-of-sight velocities with respect to the FRG of over 300 km s^{-1} are common. In fact, a few of these merging satellites have speeds of up to 800 km s^{-1}, but these are relatively few and are seen only after the FRG has grown considerably in size. It may be that a galaxy continues to exist as a dynamical unit after the friends-of-friends algorithm loses sight of it; however, visual inspection of various merging events seems to rule this out. Thus those nuclei seen by Lauer (1988) which appear to be interacting with FRG but have relative velocities of more than 400 km s^{-1} may still be in the process of merging.

The velocity dispersions of the models are generally consistent with a Gaussian as measured by statistical tests for normality (using the ROSTAT routines mentioned in § 3.4.). The possibility that there may be departures from a Gaussian before the onset of a burst of merging was investigated, but no such signal was found.

4. DISCUSSION

4.1. *Comparison with Other Simulations*

The models presented here have similarities with recent investigations by Funato, Makino, and Ebisuzaki (1993, hereafter FME) and Malumuth (1992). FME used a technique similar to ours to arrive at strikingly different results. Clusters made of 65,536 particles were run on the special-purpose GRAPE-3 computer, which integrates the softened potential N-body problem at high speed. Each galaxy was modeled by 512 or 1024 particles and a cluster model consisted of 128, 96, or 64 galaxies. There were only a few mergers in each simulation; stripping was much more important, with about half the total mass escaping from the galaxies by the end of the simulations.

The differences in evolution between the FME models and those presented here appear to stem from a different choice of initial conditions. First, the binding energy per unit mass of a FME galaxy is typically $E_g/M_g = 0.25$, whereas for the cluster as a whole $E/M = 1.6$, so the ratio of the cluster velocity dispersion to the internal galaxy dispersion is $V/V_g = 2.5$. In our models, those galaxies with $M_g > M^*$ have internal dispersions close to the cluster dispersion, making mergers more likely. The initial galaxy distribution within the cluster also plays an important role; our initial spatial configurations are quite different, as can be seen by comparing Figure 1 to FME Figure 1. FME used a Plummer model for the initial cluster density profile. The innermost regions (within half the Plummer model scale length) are quite dense, such that the intergalaxy separation is of order the galaxy size. This results in very strong tidal interactions between the galaxies in the inner region, causing them to lose mass much more quickly than galaxies in outer regions. Also, the smallest galaxies used by FME have an initial mass $M_g = M/128$, while the median initial galaxy size in our $\beta = 50$ models is slightly less than this ($M_g = M/140$). Thus in our models there are fewer large perturbers which could strip mass from a galaxy. Finally, some of the difference between our results and those of FME may simply be a matter of interpretation. FME follow galaxies even after more than 95% of the mass has been stripped (by finding which particles continue to have negative binding energy); the stripped particles form a spherically symmetric distribution in the cluster. Our galaxy-locating algorithm would probably consider the central portion of this configuration to be a large merger remnant, rather than several small galaxy nuclei moving through a background.

The initial state used here is quite similar to that of Malumuth (1992), while the numerical techniques differ (see § 1). In particular, our $\beta = 50$ models have core radii and velocity dispersions similar to Malumuth's models B, C, and H. In 10 Gyr, roughly a third of these latter models created a FRG more than twice as luminous as the initially brightest galaxy, whereas all but one of our models did so. Thus, merging and dynamical friction are apparently more efficient in our simulations. Most of Malumuth's models have a larger core radius (350 kpc) and a higher line-of-sight velocity dispersion (400 km s^{-1}); we can roughly match this by scaling the code units of mass and length to 2.5×10^{14} $M_\odot$ and $R = 2$ Mpc. This increases the size and mass of the galaxies as well, so that our M^* galaxy is twice as massive as Malumuth's; the unit of time is 2.6 Gyr. The galaxies which Malumuth identifies as cD's (those objects twice as luminous as the initial brightest member) at 10 Gyr have similar properties to our FRG at $t = 4$. They are generally within 100 kpc of the cluster center and have low peculiar velocities. Only a fifth of the models produced a galaxy large enough to fit Malumuth's definition of a cD; those galaxies that do have undergone a few more mergers and tend to be more luminous averaging $5.1L^*$) than our FRG. More than half of Malumuth's cD's have velocities less than 50 km s^{-1}; with a cluster dispersion of $\sim$400 km s^{-1}, this implies that the median Z-score would be $\sim$0.125, roughly half of our median Z. In both sets of models the FRG peculiar

velocities are less than the cluster velocity dispersion. Both sets of models also produce only a few merger remnants (since most of the mergers involve the FRG). Thus if ellipticals formed from the merging of spiral galaxies, then current clusters are not the primary environment where this took place.

4.2. *Comparison with Observations*

Our initial conditions were chosen to resemble clusters in the MKW and AWM catalogs of poor clusters that contain a D/cD galaxy. Price et al. (1991) constructed a representative sample of poor clusters and found that the MKW clusters are in fact typical of all poor clusters in their photometric and X-ray morphology. Photometric data for 16 MKW/AWM clusters were compiled by Yamagata (1986); these are V-band isophotal magnitudes with a threshold surface brightness of 24 mag arcsec^{-2}. By excluding galaxies lying outside the radial extent of the clusters as calculated by Yamagata & Maehara (1986), the apparent magnitudes of the cluster members can be found. The magnitude difference between the two brightest galaxies, ΔM_{12}, ranges from 0.1 to 1.8, with a median of 0.8 and an interquartile range of 0.4–1.45. Yamagata & Maehara (1986) also give the absolute magnitude for the brightest cluster galaxies; using the Lugger (1989) value of $M_R^* = -22.7$ and $V-R = 0.86$, the luminosity of the brightest cluster galaxy is in the range 0.8–6.0L^*, with a median of 2.5L^* and an interquartile range of 1.6–2.8L^*. Alternatively, Yamagata & Maehara (1986) calculate a best-fit Schechter function to the composite sample of 16 clusters and find $M_V^* = -21.57$, making the median FRG luminosity 3.2L^*.

Thuan & Romanishin (1981) estimate the typical luminosity of the FRG in MKW/AWM clusters at 7L^*, but they followed the galaxies out to a much lower surface brightness level (fainter than 28 mag arcsec^{-2}), thereby including more luminosity from the outer regions of the galaxy. By using the friends-of-friends algorithm we are in essence picking out particles in regions where the density is above a given limit, with this limit depending on the separation parameter as $\sim d^{-3}$. For a large merger remnant, the mass contained in the diffuse outer regions will not be counted. Thus this algorithm gives luminosities more akin to the determination of isophotal magnitudes than to methods which follow galaxy halos as far as possible.

The median ΔM_{12} of our models is below 0.1 for the first two crossing times, rising to 0.2 by $t = 3$. After $t = 3$, ΔM_{12} increases rapidly; by $t = 4$, the median is 0.5, with a range of 0.2–1.5. Thus, the models at times $t \approx 4$, or 11 Gyr, are consistent with the observed range of ΔM_{12}, although with a lower median value. At later times the FRG continues to grow rapidly, such that by $t = 5$ all the models have $\Delta M_{12} \geq 0.8$. By $t = 6$, the mean $\Delta M_{12} = 1.5$ with a range of 1–2, and at $t = 8$, $\Delta M_{12} \simeq 2$. The range of model FRG luminosities at $t \approx 4$ are also consistent with observations. At $t = 4$ the FRG luminosities are in the range 1.6–6.0L^*, with a median of 3.5L^*. As with ΔM_{12}, at earlier times the model FRG luminosities are too low ($<2L^*$ when $t < 3$) and at later times they are too large ($>6L^*$ when $t > 6$) to agree with observed clusters.

Beers & Geller (1983) have shown that D and cD galaxies in rich clusters are found near local density maxima even if they are not the brightest cluster galaxy. Similarly, Bird (1994) found 23 of 25 FRGs in rich clusters to be centrally located after accounting for substructure. The surface density of galaxies around the D/cD is cusped in rich clusters and can be fit either by a power law (Beers & Tonry 1986) or an exponential (Merrifield & Kent 1989). The creation of a large galaxy through merging and the establishment of a cusped distribution around it are accomplished in a few crossing times in our models of poor clusters. While such a cusp can partially explain the observation of multiple nuclei, more multiple nuclei are observed than can be accounted for as simply arising from a constant slope cusp (Merrifield & Kent 1989). These additional nuclei are also seen in our models, but only during a fairly narrow time range (near $t = 4$–5). The number of multiple nuclei may thus be an indicator of the dynamical age of the system. It must be old enough to have produced a FRG with a satellite population, but not so old that this population has been accreted.

For $t \geq 4$ the peculiar velocities of the FRG are too low to fit the distribution of Z-scores seen in MKW/AWM clusters. In order to have a reasonable agreement with the observations, it is necessary to have a spread of ages, with a typical age of $t \approx 3$ (8 Gyr). This indicates that some clusters are dynamically young; the clusters with the highest Z-scores would be among these.

In order for dynamical evolution to proceed at a significant rate in the 50-galaxy clusters we consider here, the galaxies must initially carry substantial dark halos. Our $\beta = 90$ models generally have too high Z-scores until $t \geq 6$ (16 Gyr). ΔM_{12} remains small and the galaxy size distribution is altered very little. The density distribution does not rise near the FRG, and the number of multiple nuclei seen in the FRG remains small throughout the evolutions, rising only near the end ($t > 7$). The $\beta = 75$ models are similar in many ways to the $\beta = 50$ ones. The main difference is that the luminosity of the FRG is lower at a given time when $\beta = 75$. It has sometimes been argued that cluster galaxies cannot have massive halos, since these would quickly be sheared off by tidal interactions (e.g., Bird, Dickey, & Salpeter 1993). While the galaxies in the $\beta = 50$ simulations do lose some halo mass over time, a substantial amount is retained. This will be examined in a later paper; preliminary tests indicate that the galaxies may retain $\sim 80\%$ of their initial mass. Thus dynamical evolution proceeds throughout the simulation, rather than being confined to early times only.

4.3. *Richer Clusters and Substructure*

The success of our models in explaining several observed cluster properties as a consequence of dynamical evolution is qualified by our consideration of only 50-galaxy models, which are appropriate to poor clusters. While Price et al. (1991) and Bahcall (1980) concluded that poor clusters have properties consistent with being an extension of rich clusters to lower masses, it may still be inappropriate to extrapolate from poor clusters to richer ones with much higher velocity dispersions, since the efficiency of merging and tidal stripping both depend on the ratio V/V_g. The galaxy orbital velocities scale as $N_g^{1/2}$ for constant β, so interaction rates will decrease with increasing N_g. Thus, our results are not in conflict with earlier findings of low dynamical evolution rates in *rich*, high-velocity-dispersion clusters (e.g., Merritt 1988; Lauer 1988; Merrifield & Kent 1991). We have carried out an $N = 10^5$ particle simulation for a 100-galaxy cluster with $\beta = 50$ that will be reported in detail elsewhere. As expected, the merging rate is reduced from the 50-galaxy models. We plan to carry out larger simulations using massively parallel processor systems.

In addition, an isolated, spherical, virialized system is an idealized situation, since roughly a third of all clusters show some evidence of substructure (the evidence is reviewed in

Richstone, Loeb, & Turner 1992). Such substructure is expected in a high- or critical-density universe as existing structures continue to grow by accretion; isolated, unperturbed structures are expected to be prevalent only in a low-density universe. Assuming that substructure is washed out on a timescale of $0.1H_0^{-1}$ (Richstone et al. 1992), then a third of present-day clusters must have accreted significant mass within this interval; this implies that a typical cluster will undergo a merger roughly every $0.3H_0^{-1} = 6$ Gyr, or 2.2 code units using the scaling of § 2. Our results are applicable even if significant growth of clusters by accretion occurs over a Hubble time, as long as they are taken to describe the internal evolution of the precursors of present-day rich clusters over the substantial time intervals between major accretion events.

The formation of cD galaxies with luminosity $\approx 10L^*$ remains an issue. We have found that even galaxies with a large velocity relative to the FRG can be consumed relatively quickly, so estimates of the current rate of cannibalism by Lauer (1988) and Merrifield & Kent (1991) may be too low. On the other hand, the cannibalism rate was less in the past than it is currently. The time needed after virialization to create a $10L^*$ cD galaxy through merging is in conflict with the younger age required by observed Z-scores. The answer may lie in the cluster formation process, particularly if clusters grow by accreting smaller groups. Small groups containing galaxies with low relative velocities will quickly merge into a single large remnant (Bode et al. 1993), which could then be introduced into a larger configuration. These subcluster merging events may, to some extent, reset the evolutionary clock, by heating up the galaxy distribution within the enlarged cluster. This would tend to slow the steady growth of the FRG through accretion of smaller satellites, since the satellite distribution would be disrupted and would require time to reform. On the other hand, the largest galaxies from each subgroup would quickly find their way to the center of the new distribution and merge together, producing an FRG growth spurt during each subcluster accretion event. The resulting FRG could have a significantly larger peculiar velocity for a period of time following such a merger than do the steadily growing FRGs in our models of isolated clusters, since it would take several crossing times for the merger cluster to relax and the new FRG to settle completely to the dynamical center.

The interaction of subclumps within a rich cluster may well be critical to the overall dynamical evolution of the cluster. Since the velocity dispersion within each subunit is lower than the overall dispersion within the cluster, dynamical evolution will proceed more rapidly before the clumps merge. Our present study may be regarded as the examination of the life of a single clump prior to its merger into a larger structure. We plan large-scale simulations of systems of merging subclusters in order to test the conjecture that dynamical evolution is only important to rich clusters that have grown by hierarchical merging.

We are grateful to David Merritt and Jeremiah Ostriker for helpful conversations, Christina Bird and Timothy Beers for kindly supplying the ROSTAT routines and poor-cluster data, and Lars Hernquist for supplying his tree code. Support for this research was provided by NSF grant AST91-19965 and NASA Astrophysics Theory Program grant NAGW-2400 to Indiana University. Support for use of National Center for Supercomputing Applications facilities was provided through grant AST91-0034N; generous amounts of computer time were also provided by University Computing Services at Indiana University. P. W. B. acknowledges the support of a Second Semester Research Fellowship from Indiana University.

REFERENCES

Albert, C. E., White, R. A., & Morgan, W. W. 1977, ApJ, 211, 309 (AWM)
Bahcall, N. A. 1980, ApJ, 238, L117
Beers, T. C., Flynn, K., & Gebhardt, K. 1990, AJ, 100, 32
Beers, T. C., & Geller, M. J. 1983. ApJ, 274, 491
Beers, T. C., Bird, C. M., Kriessler, J., & Huchra, J. P. 1994, in preparation
Beers, T. C., & Tonry, J. L. 1986, ApJ, 300, 557
Bird, C. M. 1994, AJ, in press
Bird, C. M., Dickey, J. M., & Salpeter, E. E. 1993, ApJ, 404, 81
Biviano, A., Girardi, M., Giuricin, G., Mardirossian, F., & Mezzetti, M. 1992, ApJ, 396, 35
Bode, P., Cohn, H., & Lugger, P. 1993, ApJ, 416, 17
Bothun, G. D., & Schombert, J. M. 1990, ApJ, 360, 436
Cowie, L. L., & Hu, E. M. 1986, ApJ, 305, L39
Funato, Y., Makino, J., & Ebisuzaki, T. 1993, PASJ, 45, 289 (FME)
Gebhardt, K., & Beers, T. C. 1991, ApJ, 383, 72
Girardi, M., Biviano, A., Giuricin, G., Mardirossian, F., & Mezzetti, M. 1993, ApJ, 404, 38
Hausman, M. A., & Ostriker, J. P. 1978, ApJ, 224, 320
Hernquist, L. 1987, ApJS, 64, 715
———. 1990, J. Comput. Phys., 87, 137
Hoessel, J. G., & Schneider, D. P. 1985, AJ, 90, 1648
Lauer, T. R. 1988, ApJ, 325, 49
Lugger, P. M. 1989, ApJ, 343, 572
———. 1991, in Asteroids to Quasars, ed. P. M. Lugger (Cambridge: Cambridge Univ. Press), 211
Malumuth, E. M. 1992, ApJ, 386, 420
Malumuth, E. M., & Kriss, G. A. 1986, ApJ, 308, 10
Malumuth, E. M., & Richstone, D. O. 1984, ApJ, 276, 413
Merrifield, M. R., & Kent, S. M. 1989, AJ, 98, 351
———. 1991, AJ, 101, 783
Merritt, D. 1983, ApJ, 264, 24
———. 1984, ApJ, 376, 26
———. 1988, The Minnesota Lectures on Clusters of Galaxies and Large-Scale Structure, ed. J. M. Dickey (Provo: Brigham Young Univ. Press), 175
Morgan, W. W., Kayser, S., & White, R. A. 1975, ApJ, 199, 545 (MKW)
Ostriker, J. P., & Tremaine, S. D. 1975, ApJ, 202, L113
Price, R., Burns, J. O., Duric, N., & Newberry, M. V. 1991, AJ, 102, 14
Richstone, D. 1990, in Clusters of Galaxies, ed. W. R. Oegerle, M. J. Fitchett, & L. Danly (Cambridge: Cambridge Univ. Press), 231
Richstone, D., Loeb, A., & Turner, E. L. 1992, ApJ, 393, 477
Richstone, D. O., & Malumuth, E. M. 1983, ApJ, 268, 30
Thuan, T. X., & Romanishin, W. 1981, ApJ, 248, 439
Tonry, J. L. 1985, ApJ, 291, 45
Yamagata, T. 1986, Ann. Tokyo Astron. Obs., 2d Ser., 21, 31
Yamagata, T., & Maehara, H. 1986, Publ. Astron. Soc. Japan, 38, 661
Yepes, G., Dominguez-Tenreiro, R., & Pozo-Sanz, R. del. 1991, ApJ, 373, 336
Zabludoff, A. I., Geller, M. J., Huchra, J. P., & Vogeley, M. S. 1993, AJ, 106, 1273